THE NUMERICAL INTERPRETATION OF FERMENTATION-TUBE RESULTS *

M. H. McCRADY

(From the Laboratories of the Board of Health of the Province of Quebec)

The employment of the fermentation tube in the quantitative determination of certain bacteria, particularly that of Bacillus coli and allied organisms in the sanitary analysis of water, milk, and other foods, has become so general, and the results of the fermentation test have acquired such significance, that much time and effort have been expended in the attempt to increase the precision of the method. Many problems relating to media, apparatus, and technic have been solved, and the fermentation test has become firmly established.

Little attention, however, has been given to the numerical interpretation of fermentation-tube results, altho an estimate of the number of fermenting organisms in the sample is the logical end of the examination. It may be of interest to know, for instance, that of 5 tubes, each inoculated with 0.1 c.c. of the sample, 4 show presence of the organism tested for; but much more important is the knowledge, afforded by this data, that the number of organisms in the sample is most probably about 1600 per 100 c.c. (instead of 800, as might have been inferred).

Closely associated with this question of number, is the question of precision. Comparison of results, whether one with another, or with a Standard, always involves this question of precision. For instance, suppose with one medium, 85 out of 100 tubes, each inoculated with 1 c.c. of the sample, show presence of the organism; while with another medium, only 75 out of 100 tubes are "positive." Can the difference between these two results be considered really significant?

Other questions are involved in this problem of numerical interpretation, but those of number and precision are by far the most important, and demand the first consideration.

Professor Phelps[1] has approached one phase (that of averages) of the problem, but the general problem of numerical interpretation has

* Received for publication May 8, 1915.

1. Am. Pub. Health Assn. Rep., 33, p. 9.

not, to the writer's knowledge, been discussed; and the considerable variation which obtains in laboratory practice in the use of the fermentation test, and in methods of expressing results, indicates a degree of uncertainty regarding the real significance of the results.

The present study is an attempt to define the factors which enter into this problem, and to indicate some of the more important applications of the mathematical analysis to laboratory practice.

For convenience of discussion, the problem as applied to water-analysis will be considered; and Bacillus coli, capable of fermenting the medium employed, will be the fermenting organism.

The ordinary method of expressing individual results will be followed; that is, as a fraction, the denominator denoting the number of trials, and the numerator the number of these trials which gave positive results. Thus, "2/5 in 1 c.c." means that two of five tubes, each inoculated with 1 c.c. of the sample, gave evidence of the presence of Bacillus coli.

The quantity of sample is assumed to be 100 c.c., but, as will be shown later, considerable variation from this quantity will not appreciably affect the results obtained on this assumption.

General Theory

Case 1—one dilution, one tube

The Result is Negative.—Suppose one single bacillus coli is contained in the sample of 100 c.c., and suppose one fermentation tube is inoculated with 1 c.c. of the sample.

At the moment of withdrawing this 1 c.c. of the sample, the single organism may be in this 1 c.c. or it may be in any other of the remaining 99 c.c. of the sample. There is no reason to believe that any one volume rather than any other volume will be favored with the presence of the organism. Consequently, the chances are 99 out of 100 that the organism will not be in the particular 1 c.c. withdrawn. Or, in the language of Probability, in which certainty is expressed by unity, the probability that the organism will not be contained in the 1 c.c. withdrawn is equal to 99/100, or .99. Or, to use still another mode of expression, if a great number of samples, each containing one bacillus coli, were examined in this manner, about 99 percent of the results would be negative, and the greater the number of such samples examined, the nearer would the percentage of negative results approach this figure.

Now, to illustrate the next step, suppose two coins are tossed. Of course, each coin will turn up head about half the time in the long run, and the probability of this event is then said to be ½ or 0.50 for each coin. But, according to the well-known principle of compounding separate probabilities (a principle demonstrated in any text-book of algebra), the probability of both coins turning up heads at the same throw is equal to the product of the separate probabilities, or $(½)\ (½) = ¼ = 0.25$. That is, in the long run, double heads will appear about once out of every four throws.

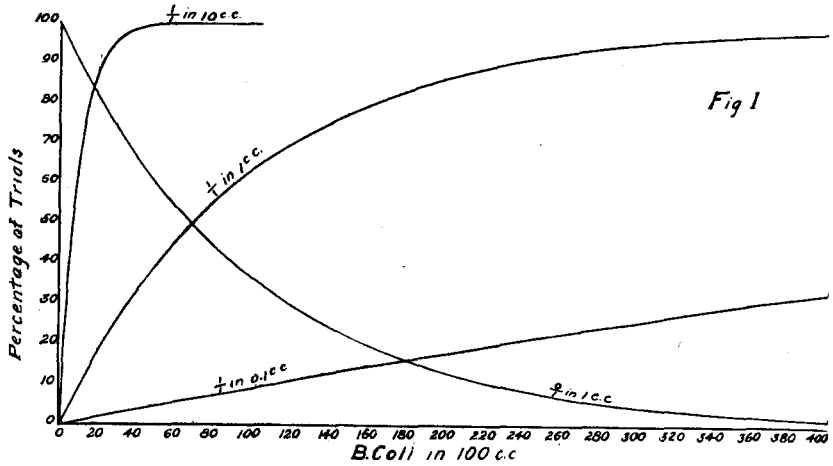

Fig. 1.—Showing the percentage of trials which will give the results indicated on the curves when certain numbers of B. coli are contained in the sample of 100 c.c.

Now suppose two B. coli are in the sample. The probability of each organism's not being contained in the 1 c.c. withdrawn for the fermentation test has been shown to be (0.99). Then by the principle just illustrated, the probability of neither organism's appearing in this 1 c.c. is equal to the product of the separate probabilities, or $(.99)\ (.99) = .9801$. And if a great number of such samples were examined, about 98.01 percent of the results would be "0/1 in 1 c.c."

In general, if V represents the number of volumes in the sample, and x the number of B. coli in the sample, and one volume is withdrawn, the probability that this volume will contain no B. coli is given by

$$\left[\frac{V-1}{V}\right]^x$$

Thus, when 1 c.c. of the sample is withdrawn for the test, V becomes 100 and the formula becomes $\left[\frac{99}{100}\right]^x$. When a 10 c.c. quan-

tity is withdrawn, V becomes 10 (there are ten 10 c.c. volumes in the sample), and the formula becomes $p9$).

By plotting values of $(.99)^x$ against given values of x, the curve "0/1 in 1 c.c.", of Fig. 1, is obtained. This curve shows, at a glance, the probability of obtaining the result "0/1 in 1 c.c.", when any given number of B. coli are contained in the sample of 100 c.c. Thus, when 230 B. coli are in the sample, the test on 1 c.c. quantities of the sample will be negative 10 percent of the time, in the long run; and the probability of obtaining a "negative" is said to be 0.10.

The Result is Positive.—Of the two possible results, "0/1 in 1 c.c.", and "1/1 in 1 c.c.", one or the other is certain to occur. Since unity represents certainty, and the probability of the one result has been found to be $(.99)^x$, the probability of the other result, "1/1 in 1 c.c.", is equal to $1-.99^x$.

In general, the same notation being used as before, the probability of the result "1/1," is given by

$$1-\left[\frac{V-1}{V}\right]^x$$

The curve for this case, when 1 c.c. is the volume withdrawn, is also shown in Fig. 1, (the curve "1/1 in 1 c.c."). The curve tends upward, indicating that the greater the number of B. coli in the sample the greater the probability of obtaining a positive result.

It is to be noticed that (for reasons to be given later) any one of the curves of Fig. 1 may be used for that corresponding to the next higher dilution, by multiplying the abscissae by ten. Thus, the probability of the result "1/1 in 1 c.c." when $x=30$, is practically identical with that for the result "1/1 in 0.1 c.c." when $x=300$. Consequently, the curve "1/1 in 1 c.c." may be used for any of the "1/1" results in the higher dilutions by multiplying the abscissae by ten, one hundred, etc.

CASE 2—ONE DILUTION, SEVERAL TUBES

It may be demonstrated (in any text-book of algebra) that if P is the probability of an event happening at one trial, the probability of its happening p times out of $p+q$ trials is given by

$$\frac{(p+q)!}{p!\,q!}(P)^p(1-P)^q$$

It follows, then, that if 2 tubes are each inoculated with 1 c.c. of the sample, the probability of obtaining the result "1/2 in 1 c.c." is given

by $1.2/1.1$ $(1-.99^x)$ $(.99^x) = 2(1-.99^x)$ $(.99^x)$ for here $P = 1-.99$; the probability of obtaining a positive result at one trial. And $p = 1$, for there is to be one positive result, and $q = 1$, for there are to be $p + q = 2$, trials.

The curve for this result, "1/2 in 1 c.c.", is shown in Fig. 3, at the right of the sheet. It indicates, for instance, that if the sample

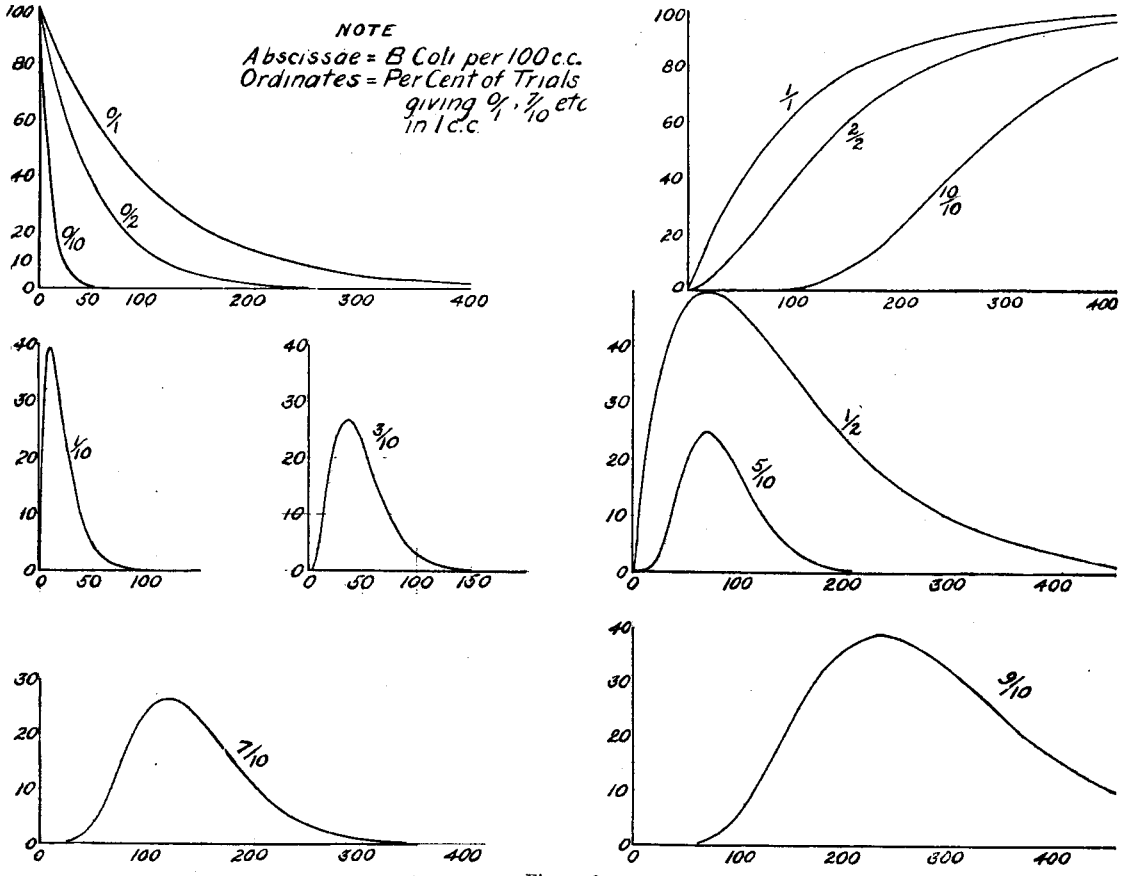

Figure 3

contains 230 B. coli, the result "1/2 in 1 c.c." will be obtained about 17.5 percent of the time, in the long run. From the other curves of Fig. 3 the result "0/2 in 1 c.c." will occur about 0.50 percent of the time and the result "2/2 in 1 c.c." about 81.0 percent of the time, respectively. Since one or the other of these results is certain to occur, the sum of these probabilities should equal unity, or 100 percent.

It is to be noticed that the number of B. coli which will give the result "1/2 in 1 c.c." with the greatest frequency is 69, and not 50, as might have been inferred.

Fig. 3 shows curves for many results of this case, comparison of which may prove of interest.

CASE 3—SEVERAL DILUTIONS, ONE TUBE AT EACH DILUTION

This very common case in which 10 c.c. of the sample are inoculated into one tube, 1 c.c. into another tube, 0.1, c.c. into another tube, and so on, submits very readily to calculation.

Thus, the probability of the result, "++—" (viz., 1/1 in 10 c.c., 1/1 in 1 c.c., 0/1 in 1 c.c.) is given by the product of the separate probabilities corresponding to the separate parts of the compound result, or

$$(1-.9^x)\ (1-.99^x)\ (.999^x)$$

The curve for this result, together with that for the result +— — and for the "anomaly," (+—+), is shown in Fig. 2.

To illustrate the use of these curves, suppose a number of samples each containing 100 B. coli were examined by this system of one tube at each of several dilutions, say three, 10 c.c., 1 c.c., and 0.1 c.c. From the curves it is seen that the results obtained would be about as follows: (+——) 32.5 percent of the time; (++—) 58.5 percent of the time; and the anomaly (+—+) 3.5 percent of the time. The remaining 5.5 percent of the results would be distributed among the other possible results: (+++), (—++), and (———).

It is to be noticed (for reasons to be mentioned later) that any one of these curves may be used for that of the next higher combination, by multiplying the abscissae by ten. Thus, the ordinate for $x=20$, in the +——) curve is practically identical with that for $x=200$, in the (++—) curve. (For large values of x the curve (+——), with abscissae multiplied by ten would correspond more nearly to the curve (++——), with formula $(1-.9^x)$ $(1-.99^x)$ $(.999^x)$ $(.9999^x)$, but for smaller values of x, the last factor, $(.9999^x)$ is practically equal to unity and may be disregarded.)

The highest ordinates of the curves for this case correspond to the numbers of B. coli 23, 230, 2300, etc., and in the case of the anomalies, to 9, 90, etc., approximately.

CASE 4—SEVERAL DILUTIONS, SEVERAL TUBES AT EACH DILUTION

This case, altho more complicated, may be readily analyzed. The formulae are made up by simply compounding the separate probabilities of the separate parts of the compound result, just as was done in Case 3.

Thus, the probability of obtaining the result "2/2 in 10 c.c., 3/10 in 1 c.c., 0/10 in 0.1 c.c." is given by:

$$\left[(1-.9^x)^2\right]\left[120\,(1-.99^x)^3\,(.99^x)^7\right]\left[(.999^x)^{10}\right]$$

Such formulae are very easily reduced with the aid of a table of logarithms; and, by plotting a few of the ordinates, the general shape of the curve for any result may be readily determined.

This Case 4, of course, includes the cases already discussed.

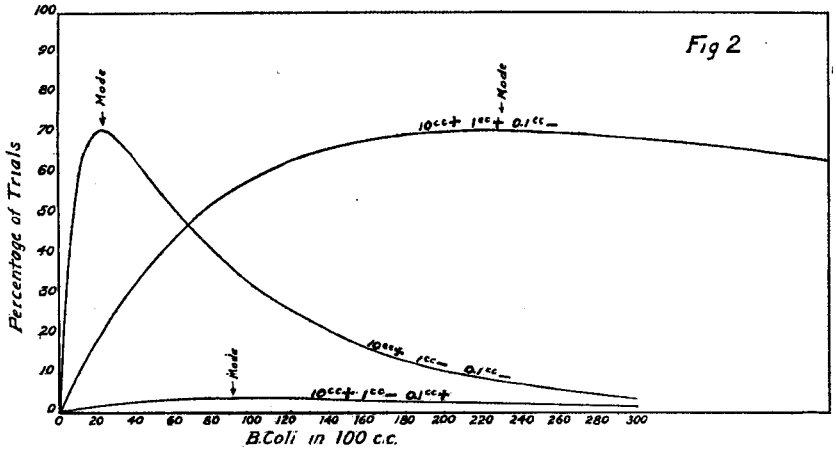

Fig. 2.—Showing the percentage of trials which will give the results indicated on the curves when certain numbers of B. coli are contained in the sample of 100 c.c.

ASSUMPTIONS INVOLVED IN THESE FORMULAE

Quantity of sample.—The formulae given assume the quantity of sample to be 100 c.c. But in the expression $\left[\frac{V-1}{V}\right]^x$, the basic factor of all the formulae, even though V varies, so long as x varies proportionately, the value of the factor remains practically unchanged. Thus, suppose one sample of 100 c.c. containing 50 B. coli, and another sample of 80 c.c. containing 40 B. coli. The factor becomes, for these two cases, $(.99)^{50}=.6050$ and $\left[\frac{79}{80}\right]^{40}=.6046$, or practically identical values.

In fact, a great variation, in excess, from 100 c.c., will have little effect on this factor. Thus, suppose one sample of 100 c.c. containing 100 B. coli, and another of 10,000 c.c. containing 10,000 B. coli. The factor becomes, for these two cases, $(.99)^{100}=.366$ and $(.9999)^{10,000}=.368$, respectively. This fact accounts for the facility with which one curve of Fig. 1, Fig. 2, or Fig. 3 may be employed for the curve of the next higher order, as has already been noted under Case 1 and Case 3.

Consequently, the variation in quantity of sample from 100 c.c. obtaining in ordinary laboratory practice will have little effect upon these formulae. To be sure, in special cases the variation may have to be allowed for, but in general the formulae may be applied directly to laboratory results.

Replacement.—When more than one volume is to be drawn from the sample, these formulae demand that for each draw the initial conditions must be the same. That is, after the first volume has been drawn, this volume, together with its contained B. coli, must be replaced in the sample-bottle before drawing the next volume. Such a procedure is obviously impossible in practice.

But perhaps, when the first volume has been drawn, it may be assumed that a proportionate number of the B. coli have also been drawn in this volume. If so, since both V and x have varied proportionately, the value of the general factor $\left\{\frac{V-1}{V}\right\}^x$ has remained practically unchanged.

But even if this assumption is not justified, calculation will show that the error due to this non-replacement is, in general, negligible. (Some experiments, to be described later, show quite clearly the negligible character of this error.)

For ordinary practical work, even when small numbers of B. coli are in the sample, these formulae may be employed. But in certain special work, and wherever extreme accuracy is desired, it may be necessary to construct a new set of formulae which will allow for non-replacement, as is done later in the discussion of the recently established United States Treasury Standard for waters supplied by common carriers.

APPLICATIONS

THE "MOST PROBABLE NUMBER"

Its Significance.—In all sampling, whether of population, of chemical substance, or of bacteria, a certain line of reasoning is employed which is not always recognized.

Consider the plate method of examining a water for bacteria. Suppose 1 c.c. of the sample is used, and suppose the plate count is found to be 25. The bacterial content of the sample is then recorded as 25 per cubic centimeter. But it is recognized that any number of bacteria in excess of 24 may have been contained in that sample. In other words, the record "25 per cubic centimeter" may be in error by almost any amount. Wherein, then, lies the justification for choosing this particular number, 25, to represent the average number of bacteria per cubic centimeter of the sample? It may be said that "common sense" justifies the choice. But is the argument so simple? in fact, in the exactly analogous case of the fermentation-tube result, "common sense" does not lead one very far.

The real justification is to be found in the following line of reasoning:

If the number of bacteria in the sample of 100 c.c. is 2500, the probability of obtaining the plate count 25 from 1 c.c. of the sample is equal to P_1 (given, as may be shown,* by the twenty-sixth term in the expansion of the binomial $\left\{\frac{99+1}{100}\right\}^{2500}$).

If the number of bacteria is 2501, the probability is P_2.

If the number of bacteria is 2502, the probability is P_3, and so on.

But of all these various probabilities, it is found that P_1, corresponding to 2500 bacteria, is the greatest.

Now "common sense" may be enlisted in the form of this principle[2]: "Of all those values of an unknown quantity which, before the occurrence of a certain event, were equally probable, that one is after the event the most probable which, before the event, assigned to it the greatest probability."

Therefore, since 2500 is that value of the unknown quantity which assigns the greatest probability to the occurrence of the event (the

* The probability of drawing exactly k organisms in the one volume drawn from the sample of V volumes, when the sample contains x organisms, is given by the $(k+1)$th term in the expansion of the Binomial $\left\{\frac{(V-1)+1}{V}\right\}^{x}$.)

2. Johnson: Theory of Errors and Method of Least Squares, 1905, p. 18.

plate-count 25), 2500 is the most probable number of bacteria in the sample, and should be so recorded.

Now notice the exactly parallel reasoning in the case of the fermentation-tube result. Take the result "9/10 in 1 c.c.":

If the number of B. coli in the sample was 229, the probability of obtaining this result is equal to P^1 (given by $10\ (1-.99^{229})^9\ (.99^{229})$).

If the number of B. coli was 230, the probability is P_2.

If the number of B. coli was 231, the probability is P_3, and so on.

Now of all these probabilities, the greatest is found to be that corresponding to the number 229.

Therefore, by the principle just enunciated, 229 is the most probable number of B. coli in the sample.

It is evident, then, that these two "most probable numbers," 2500 bacteria, and 229 B. coli, have exactly the same status in their respective domains. Each is obtained in exactly the same manner as is the other. The fermentation-tube result, "9/10 in 1 c.c.," means 229 B. coli per 100 c.c. just as the plate-count "25" means 25 bacteria per cubic centimeter.

Now either of these "most probable numbers" may be in error. In the case of the plate count, for instance, the number of bacteria in the sample may be 10,000 instead of 2500. But in the long run of samples of various waters, the application of the line of reasoning described will lead to a series of "most probable numbers" which, on the whole, will strike closer to the truth than will any other series of numbers obtained in any other way. Each "most probable number" represents the one best guess afforded by the data of the analysis, no matter whether the result takes the form of a plate count, a fermentation-tube result, or a chemical proportion.

Calculation of the Most Probable Number.—In the case of the plate count, it happens that this count may be converted directly into its corresponding most probable number of bacteria. The most probable number is, so to speak, a linear function of the plate count (the plate count being usually simply multiplied by one).

But in the case of the fermentation-tube result, the most probable number of B. coli is a logarithmic function of this result, and recourse must be had to calculation to obtain the most probable number.

Of course, one method of obtaining the most probable number corresponding to the result, say "5/10 in 1 c.c.," is to plot the curve for

this result and then pick out the highest ordinate and take the corresponding number of B. coli.

But ordinarily, especially for compound results, such a method is too laborious, and some short-cut by calculation is desirable.

It may be shown that, given the result "$\frac{p}{p+q}$ in 1 volume," the corresponding most probable number is given by the solution for x of the equation

$$1-\left[\frac{V-1}{V}\right]^x=\frac{p}{p+q}.$$

Thus, for the result "5/10 in 1 c.c.," the most probable number is given by solution of the equation $1-.99^x=5/10$. The equation being solved, $x=69$. This is the most probable number of B. coli in the sample, per 100 c.c.

(It is to be noticed that the curve "1/1 in 1 c.c.", of Fig. 1, is a graph of this general equation with V equal to 100.)

From the form of the equation, it is evident that multiple results, such as "1/2 in 1 c.c." and "5/10 in 1 c.c.," correspond to the same most probable number of B. coli.

Application of the formula gives some interesting information. Thus, the result "9/10 in 1 c.c." means, not 90 B. coli per 100 c.c., but 229 B. coli per 100 c.c., altho the result "1/10 in 1 c.c.," means 10 B. coli, as might have been presumed.

For compound results, a more complicated formula must be employed. This equation may be built up as follows:

(1) Suppose the result is "$\frac{p}{p+q}$ in 10 c.c." The equation becomes

$$(p+q)\ (\log .9)=\frac{p\ (\log .9)}{1-.9^x}$$

(These equations are obtained by differentiating for a maximum the equation for the probability of the result.)

(2) Suppose the result is $\frac{p}{p+q}$ in 10 c.c., $\frac{r}{r+s}$ in 1 c.c." The equation becomes

$$(p+q)\ (\log .9)+(r+s)\ (\log .99)=\frac{p\ (\log .9)}{1-.9^x}+\frac{r\ (\log .99)}{1-.99^x}.$$

The manner in which corresponding terms are added to each side of the first equation to build up the second equation is very apparent. To take a concrete example, suppose the result is "½ in 10 c.c., 3/10 in 1 c.c., 0/10 in 0.1 c.c." The equation becomes

$$2\ \log .9+10\ \log .99+10\ \log .999=\frac{\log .9}{1-.9^x}+\frac{3\ \log .99}{1-.99^x}+0.$$

and solution of this equation for x gives $x = 17$, to the nearest unit, and the most probable number of B. coli in the sample is 17 per 100 c.c.

These equations for compound results must be solved by "trial and error," but the work proceeds very rapidly, and is not so laborious as it may appear.

For the particular system employed by each laboratory, it would be advisable to calculate, once for all, the most probable numbers corresponding to the results which this system may give. A rapid review of past results will indicate the range of practically possible combinations for which the most probable numbers should be calculated. For instance, in Table 1 are given the most probable numbers corresponding to most of the practically possible results which may occur from the system: Two tubes at 10 c.c., ten tubes at 1 c.c., and ten tubes at 0.1 c.c. Of course, still other combinations, not included in this table, are sure to occur sooner or later, and must be given their most probable numbers; but such other combinations will not occur often.

Every Result Must Be Given Its Interpretation.—Just as every plate-count is given its most probable interpretation, so should every fermentation-tube result be given its most probable interpretation. Just as it may seem absurd to think of the logarithm of the plate count when considering the number of bacteria in the sample, so should it seem absurd to think of the result "2/3 in 1 c.c." when considering the number of B. coli in the sample. This result means, so far as the analytical result can signify, 109 B. coli per 100 c.c. and this is the record of the analysis which must be considered.

By the methods which have been described, every result, whether simple or compound (except the single result N/N, such as "3/3 in 1 c.c.") may be interpreted. Compound results, particularly, require such interpretation, for it is often quite difficult to guess the significance of such results.

Moreover, the odd result, such as the "anomaly" in the system of one tube at each dilution, must not be generally regarded with suspicion. In fact every possible result, every possible combination, is sure to be obtained sometime. Overgrowths and other cultural difficulties may be responsible for some odd results, but, in general, the benefit of the doubt must be given the chance distribution, for it is the one cause of such results which is known to be continually operating.

Calculation will often show whether the proportion of odd results obtained on a series of samples is to be suspected. Thus, the curve of Fig. 2 shows that the anomalies will occur about 3 or 4 percent of the time, on the average; therefore, a proportion of anomalies greatly in excess of this may be regarded with suspicion.

TABLE 1

GIVING THE "MOST PROBABLE NUMBERS" OF B. COLI PER 100 C.C. OF SAMPLE CORRESPONDING TO VARIOUS FERMENTATION-TUBE RESULTS

Using Two Dilutions			Using Three Dilutions															
10 c.c.	1 c.c.	Number	10 c.c.	1 c.c.	0.1 c.c.	Number	10 c.c.	1 c.c.	0.1 c.c.	Number	10 c.c.	1 c.c.	0.1 c.c.	Number	10 c.c.	1 c.c.	0.1 c.c.	Number
0/2	0/10	0	0/2	0/10	0/10	0	1/2	4/10	0/10	22	2/2	4/10	4/10	94	2/2	8/10	5/10	245
0/2	1/10	3	0/2	0/10	1/10	3	1/2	4/10	1/10	27	2/2	4/10	5/10	105	2/2	8/10	6/10	275
0/2	2/10	7					1/2	4/10	2/10	32					2/2	8/10	7/10	310
0/2	3/10	10	0/2	1/10	0/10	3	1/2	4/10	3/10	37	2/2	5/10	0/10	61				
0/2	4/10	14	0/2	1/10	1/10	6					2/2	5/10	1/10	72	2/2	9/10	0/10	170
							2/2	0/10	0/10	10	2/2	5/10	2/10	86	2/2	9/10	1/10	195
1/2	0/10	4	0/2	2/10	0/10	6	2/2	0/10	1/10	16	2/2	5/10	3/10	100	2/2	9/10	2/10	225
1/2	1/10	8	0/2	2/10	1/10	10	2/2	0/10	2/10	23	2/2	5/10	4/10	115	2/2	9/10	3/10	260
1/2	2/10	13					2/2	0/10	3/10	30	2/2	5/10	5/10	130	2/2	9/10	4/10	300
1/2	3/10	17	0/2	3/10	0/10	10					2/2	5/10	6/10	145	2/2	9/10	5/10	345
1/2	4/10	23					2/2	1/10	0/10	17					2/2	9/10	6/10	395
1/2	5/10	28	0/2	4/10	0/10	13	2/2	1/10	1/10	23	2/2	6/10	0/10	79	2/2	9/10	7/10	455
1/2	6/10	35					2/2	1/10	2/10	31	2/2	6/10	1/10	93	2/2	9/10	8/10	525
			1/2	0/10	0/10	4	2/2	1/10	3/10	41	2/2	6/10	2/10	110				
2/2	0/10	11	1/2	0/10	1/10	8	2/2	1/10	4/10	49	2/2	6/10	3/10	125	2/2	10/10	0/10	240
2/2	1/10	18									2/2	6/10	4/10	140	2/2	10/10	1/10	285
2/2	2/10	27	1/2	1/10	0/10	8	2/2	2/10	0/10	24	2/2	6/10	5/10	155	2/2	10/10	2/10	345
2/2	3/10	37	1/2	1/10	1/10	12	2/2	2/10	1/10	33	2/2	6/10	6/10	170	2/2	10/10	3/10	430
2/2	4/10	52	1/2	1/10	2/10	16	2/2	2/10	2/10	42					2/2	10/10	4/10	550
2/2	5/10	69	1/2	1/10	3/10	20	2/2	2/10	3/10	52	2/2	7/10	0/10	100	2/2	10/10	5/10	720
2/2	6/10	91					2/2	2/10	4/10	62	2/2	7/10	1/10	120	2/2	10/10	6/10	920
2/2	7/10	120	1/2	2/10	0/10	12					2/2	7/10	2/10	135	2/2	10/10	7/10	1,200
2/2	8/10	160	1/2	2/10	1/10	16	2/2	3/10	0/10	34	2/2	7/10	3/10	155	2/2	10/10	8/10	1,600
2/2	9/10	230	1/2	2/10	2/10	21	2/2	3/10	1/10	44	2/2	7/10	4/10	175	2/2	10/10	9/10	2,300
			1/2	2/10	3/10	26	2/2	3/10	2/10	54	2/2	7/10	5/10	195				
							2/2	3/10	3/10	65	2/2	7/10	6/10	215				
			1/2	3/10	0/10	17	2/2	3/10	4/10	77	2/2	7/10	7/10	235				
			1/2	3/10	1/10	21	2/2	3/10	5/10	88								
			1/2	3/10	2/10	26					2/2	8/10	0/10	130				
			1/2	3/10	3/10	31	2/2	4/10	0/10	46	2/2	8/10	1/10	150				
							2/2	4/10	1/10	57	2/2	8/10	2/10	170				
							2/2	4/10	2/10	69	2/2	8/10	3/10	195				
							2/2	4/10	3/10	81	2/2	8/10	4/10	220				

These "most probable numbers," when less than 100, are correct to the nearest unit. When over 100, they are correct to the nearest 5 units.

But unless there is overwhelming evidence in favor of rejecting any particular result, it should be recorded, and given its most probable interpretation.

THE PRECISION OF A RESULT

Obvious Limits.—A notion of the precision of a result may be obtained simply by inspection of the curve for this result. Thus, suppose the result "5/10 in 1 c.c." Inspection of the corresponding curve

of Fig. 3 shows that the practically possible numbers of B. coli have a range from about 15 to about 200 per 100 c.c.

Moreover, by an extension of the reasoning, already outlined, for the determination of the significance of the "most probable number," it may be shown that the relative heights of the ordinates of the curve give the relative probabilities that the corresponding abscissae were responsible for the result. Consequently, the general shape of the curve indicates roughly the degree of confidence which may be assigned to the inclusion of x within certain limits. Thus, with the result "1/10 in 1 c.c.," it is quite certain that the number of B. coli lies within the range 0 to 50. But with the result "9/10 in 1 c.c.", it is far from certain that the number lies between even 200 and 300.

Again, consider the multiple results "1/2 in 1 c.c." and "5/10 in 1 c.c." These results both correspond to the same "most probable number," but the precision of the one is quite different from that of the other. And inspection of the corresponding curves gives some indication of the extent to which the respective precisions differ.

(A more elaborate and exact method of determining the precision, based on this general principle, is given further on.)

Significance of Certain Results.—Suppose a stream sampled at two points. Suppose the result on the upper sample to be "0/1 in 1 c.c." and the result on the lower sample to be "1/1 in 1 c.c." Can the difference between these two results be considered significant? Obviously, from the manner in which the two corresponding curves of Fig. 1 overlap, the presumption in favor of the greater B. coli content obtaining at the lower point is very slight indeed.

Suppose the stream contained uniformly 69 B. coli 100 c.c. The probability of obtaining the result "0/1 in 1 c.c." on the upper sample and the result "1/1 in 1 c.c." on the lower sample, is given by the product of the two separate probabilities, or

$$(.99^{69})\ (1 - .99^{69}) = (.5)\ (.5) = .25$$

This same probability holds for obtaining the result "0/1 in 1 c.c." on the lower sample and the result "1/1 in 1 c.c." on the upper sample. The sum of these two probabilities is 0.50, and therefore, half the time, when sampling at these two points, the result on one sample will be positive and on the other sample negative. One quarter of the time both will be positive, and one quarter of the time both will be negative. (It will be noticed that this case is exactly analogous to that of tossing two coins, previously discussed.)

Again, suppose the upper stream contained 160 B. coli per 100 c.c., and the lower stream only 90. The probability of obtaining the result "0/1 in 1 c.c." at the upper point, and "1/1 in 1 c.c." at the lower point, is given by

$$(.99^{160})\ (1-.99^{90}) = (.2)\ (.6) = .12$$

Consequently, about once every eight times that these conditions obtain, the inference would be that the B. coli content of the lower stream was greater than that of the upper stream, and this despite the fact that the upper stream contained nearly twice the number of B. coli contained in the lower stream.

These examples show very clearly the necessity of an exhaustive examination of every sample (the employment of many trials), as well as of frequent sampling, especially when results are to be compared. They also indicate how, by a few simple calculations, some notion of the precision of a result may be obtained. In the instances given, such calculations have revealed beyond question the utter untrustworthiness of a system of one tube at each dilution in work of a closely comparative nature.

Precision by Bayes' Theorem.—By a principle of Probability known as Bayes' Theorem,[3] it may be demonstrated that, given the result "1/2 in 1 c.c.", the probability that the sample contained less than k B. coli per 100 c.c. is given by dividing the sum of the ordinates from $x=0$ to $x=k-1$, of the curve for this result, by the sum of all the ordinates of the curve. This principle is an elaboration of that previously mentioned, namely, that the relative heights of the ordinates give the relative probabilities that the corresponding numbers of B. coli were responsible for the result obtained.

The summation of the ordinates may be effected as follows: For any result of the form "$\frac{p}{p+q}$ in 1 c.c.", the sum of all the ordinates of the curve is given by

$$\frac{(p+q)!}{p!\,q!}\left\{\frac{1}{1-.99^{q}} - p\,\frac{1}{1-.99^{(1+q)}} + \frac{p(p-1)}{2!}\,\frac{1}{1-.99^{(2+q)}} \cdots\cdots \text{to } p+1 \text{ terms}\right\}$$

And the sum of the ordinates from $x=k$ to $x=\infty$, inclusive, is given by

$$\frac{(p+q)!}{p!\,q!}\left\{\frac{.99^{kq}}{1-.99^{q}} - p\,\frac{.99^{k(1+q)}}{1-.99^{(1+q)}} + \frac{p(p-1)}{2!}\,\frac{.99^{k(2+q)}}{1-.99^{(2+q)}} \cdots\cdots \text{to } p+1 \text{ terms}\right\}$$

3. Todhunter: History of the Theory of Probability, 1865, p. 294; Poincaré: Calcul des Probabilités, 1912, p. 154.

As an example, take the result "1/2 in 1 c.c." Here $p=1$, $q=1$, and the sum of all the ordinates is

$$2\left[\frac{1}{1-.99}-\frac{1}{1-.99^2}\right]=99.5025$$

and the sum of the ordinates from say $x=300$ to $x=$infinity, inclusive, is

$$2\left[\frac{.99^{200}}{1-.99}-\frac{.99^{600}}{1-.99^2}\right]=9.566$$

The sum of the ordinates to the left of $x=300$ is then

$$99.503 - 9.566 = 89.937$$

Therefore the probability that the number of B. coli is less than 300 per 100 c.c., is $\frac{89.937}{99.503}$, or the odds are 89.937 to 9.566, or about nine to one, that the number of B. coli is less than 300 per 100 c.c. rather than 300 or more.

For the 0/N curves, these formulae reduce very simply. Each ordinate becomes the probability that the number of B. coli in the sample is equal to, or greater than, the corresponding abscissae. Thus with the result "0/2 in 1 c.c.", its curve (Fig. 3) shows that the ordinate corresponding to $x = 150$ is about 0.05, and consequently the odds are about 95 to 5 that the number of B. coli in the sample is less than 150 per 100 c.c.; or, in other words, in analyzing a large number of samples of various waters, about 95 percent of those samples which give the result "0/2 in 1 c.c." will have contained less than 150 B. coli per 100 c.c.

There is a certain assumption involved in this principle which must be recognized. This assumption is that all numbers of B. coli are equally probable; that is, that in the long run of samples, one number of B. coli will appear about as often as any other number.

It is to be noticed that this assumption is the same, in kind, as is involved in the determination of the "most probable number." There, it is assumed that the "most probable number," so determined, is as probable as any other number. To illustrate, suppose a sample of sewage gave a plate-count of 25 per c.c. This count would immediately be regarded with suspicion, because the analyst's experience tells him that the bacterial content of 25 per c.c. is not as probable as many other bacterial contents. Consequently, he does not trust the "most probable number." But whenever the analytical result is accepted, this attendant assumption is accepted also.

But in the application of the principle under consideration, the assumption is bolder. It demands that all the various numbers of B. coli shall be equally probable.

(Before proceeding further, it must be noted that, altho the curves theoretically extend to infinity on the right, those ordinates which are great enough to affect the summations all lie within a fairly short range. Thus, with the curve "0/2 in 1 c.c.", the ordinates beyond x=300 are practically equal to zero, and may be neglected; so that, for this result, Bayes' Theorem demands only that numbers of B. coli from 0 to 300 shall be equally probable.)

But considerable justification may be found for this apparently bold assumption. Experience teaches that, in general, one number of B. coli does occur about as often as another. To be sure the filter operator will have good reason to doubt the applicability of the assumption to the result "2/3 in 1 c.c.", on the effluent of his filter, for experience tells him that the larger values of x at the right of the curve for this result are not as likely to occur as are the smaller values at the left of the curve. But, on the other hand, the State or Provincial analyst, who is examining a great number of samples of various waters, may perhaps feel justified in assuming that, roughly, over a rather wide range, one number of B. coli is turning up about as often as any other number.

"When the probability is unknown," says Laplace,[4] "we may equally suppose it to have any value between zero and unit." Again, according to Edgeworth, quoted by Pearson,[5] "The assumption that any probability constant about which we know nothing in particular is as likely to have one value as another, is grounded upon the rough but solid experience that such constants do, as a matter of fact, as often have one value as another."

In any event, despite the difficulty involved in the assumption upon which Bayes' Theorem is based, the application of the theorem to fermentation results appears to offer some interesting possibilities. It is the nearest approach that may be made to a concrete notion of the precision of a result. The "most probable number" is extremely important, but it is known to be subject to error. The shape of the curve gives some idea of the facility of this error. But calculation of the precision affords, wherever the assumption involved is even approximately justified, a degree of assurance which permits of a fairly definite

4. Introduction, Théorie Analytique des Probabilités, 1814.
5. Grammar of Science, 1911, p. 146.

conclusion regarding the B. coli content of the sample. To say the least, there is a certain amount of satisfaction to be derived from thus placing, so to speak, a limit on the error.

It is to be noticed that this method of calculating the precision of a result is analogous to that employed in many other branches of science, such as Physics, Astronomy, Precise Surveying, in fact, wherever the Method of Least Squares is employed.

COMPARISON OF THE PLATE METHOD AND THE FERMENTATION-TUBE METHOD OF ANALYSIS

Before the more recent fermentation-tube methods of analysis came into general use, the litmus lactose agar plate method of estimating B. coli was rather widely employed. And there are, perhaps, not a few workers who are still inclined to favor the plate method on the grounds that greater precision characterizes the results obtained by this method.

Suppose there were available some plate method comparable with the fermentation-tube method, in respect of faculty of growth of B. coli.

Let a number of samples, each containing 100 B. coli, be examined by both methods, as follows: (1) in each of two plates, 1 c.c. of the sample; (2) in each of ten fermentation-tubes, 1 c.c. of sample.

The percentage of times that 0 B. coli, 1 B. coli, 2 B. coli, etc., will appear on the two plates, is given by the first, second, third, etc., terms of the expansion of the binomial $\left[\frac{49+1}{50}\right]^{100}$.

The percentage of times that the results 1/10, 2/10, etc., will occur with the tubes may be calculated by the methods already given.

When 2 B. coli appear on the two plates, the most probable number in the sample is, of course, 100 per 100 c.c.; when 4, 200 per 100 c.c., etc.

The most probable numbers corresponding to the fermentation-tube results may be calculated by the methods already given.

In Table 2 are shown these calculations: the percentage of times that the various results will occur, in the long run, together with the most probable numbers of B. coli corresponding to these results.

Comparison of these figures shows clearly that the advantage in precision is with the tube method of analysis, when ten tubes are used

against two plates; for it will indicate numbers of B. coli which are closer to the true number, 100, oftener than will the plate method.

With a much greater number of B. coli in the samples, however, the system of two plates might give the better results, for then the necessity of using dilutions would decrease the precision of the tube method. But in practice, the numerous other bacteria usually associated with large numbers of B. coli tend to over-grow and obscure the latter, to the end that a limit is defined, beyond which quantities of sample less than 1 c.c. must be plated, with a consequent decrease in the precision of the plate method.

TABLE 2

A COMPARISON OF THE MOST PROBABLE NUMBERS OBTAINED ON THE ONE HAND BY THE PLATE METHOD AND ON THE OTHER BY THE FERMENTATION-TUBE METHOD OF ESTIMATING B. COLI

Plate Method			Fermentation-Tube Method		
Result (Count) (Sum of 2 Plates)	Percentage of Time	Most Probable Number B. Coli	Result	Percentage of Time	Most Probable Number B. Coli
0.....	13.26	0	0/10.....	0.00+	0
1.....	27.98	50	1/10.....	0.07	10
2.....	27.33	100	2/10.....	0.58	22
3.....	18.23	150	3/10.....	2.69	35
4.....	9.00	200	4/10.....	8.16	51
5.....	3.53	250	5/10.....	16.95	69
6.....	1.14	300	6/10.....	24.47	91
Over 6.....	0.57	Over 300	7/10.....	24.22	120
			8/10.....	15.73	160
			9/10.....	6.05	229
			10/10.....	1.05	Over 229

In any event, the tube method of analysis is evidently quite as precise as other methods in general use, especially when ten tubes are employed. (A convenient device by means of which ten tubes may be employed as a unit is described further on.)

EXPRESSION OF RESULTS

Averages.—Because of the fact that the number of B. coli in a sample is a logarithmic function of the frequency of their appearance in the volume drawn, the problem of averaging results is very difficult of solution. It is somewhat analogous to that of finding the average of a series of numbers when the average of their logarithms is given.

Take the two results "1/10 in 1 c.c.", and "9/10 in 1 c.c.". The average of the results is "5/10 in 1 c.c.", the most probable number corresponding to which, is 69 B. coli per 100 c.c. But the most prob-

able number corresponding to the first result, "1/10 in 1 c.c.", is 10 B. coli, and that corresponding to the other results is about 230 B. coli. The average of these two most probable numbers is 120, a figure quite different from the figure 69.

Indeed, it is difficult to find any form of average which will answer all requirements. After considerable work on the problem, the writer is inclined to favor the average of the "most probable numbers" (just as is done with plate counts) for all results involving the use of more than one tube at each dilution. For the latter case, where only one tube at each dilution is used, this method of averaging "most probable number" appears to break down, but the method offered by Professor Phelps[6] seems to give fairly good averages. This method assumes the frequency of the appearance of the organism in the volume tested to vary directly with the numbers of organisms in the sample, an assumption which is not in accordance with theory. Yet, so long as anomalies are excluded from the calculation, the average by this method will not often be greatly in error, except when very short series are averaged. Calculation indicates that averages (by Phelps' method) of long series of results will be rather uniformly in excess of the true average, while averages of short series may be either in excess or defect of the true average. (Further study on this problem of averaging is in progress.)

Expression of Results.—This brief discussion shows clearly the futility of expressing results by the form of average, "B. coli present in 1 c.c. 75 percent of the time." Such an expression conveys practically no information, and is quite worthless for purposes of comparison.

On the whole, since averages are so untrustworthy, it appears that the best method of expressing results is that of giving each result in full, together with the corresponding most probable number, thus:

10 c.c.	1 c.c.	0.1 c.c.	B. coli per 100 c.c.
½	2/10	1/10	16
+	+	—	230
—	+	—	9

Such expressions give all the data of the result, so that one may determine the precision of the result, and also express the result as a concrete number of B. coli. The advantage, to the uninitiated, in having these concrete numbers given is obvious. What analyst has not labored with officials in an attempt to explain the fermentation-tube

6. Am. Pub. Health Assn. Rep., 1907, 33, p. 9.

result, and usually with little success? They can easily comprehend the meaning of the plate-count, but the involved, compound, fermentation-tube result, expressed as such, is not so readily understood, and indeed often appears to contain contradictions.

THE U. S. TREASURY STANDARD[7]

The recently established U. S. Treasury Standard for waters supplied by common carriers, demands that not more than one of five separate 10 c.c. quantities of the sample shall afford evidence of the presence of B. coli. It may be of interest to discuss the probabilities for this case.

Here the B. coli content of the samples will be rather small, as a rule, and moreover the sample is considerably depleted during the examination (usually more than half of the sample being withdrawn for the fermentation tests). Consequently, the assumptions involved in the general formulae already given do not hold, and a new set of formulae must be constructed for this particular case.

It may be shown that, for this case, the following formulae hold:

Probability of the result "0/5 in 10 c.c." is $(.5)^x$.
Probability of the result "1/5 in 10 c.c." is $5(.6^x - .5^x)$.
Probability of the result "2/5 in 10 c.c." is $10[.7^x - 2(.6^x) + .5^x]$.

These formulae allow for non-replacement, and require only that the quantity of sample be close to 100 c.c.

Some calculations, by means of these formulae, indicating the percentage of various results which will occur when various numbers of B. coli are contained in the sample, are shown below:

No. of B. Coli In Sample x	Pct. of Time that "0/5 in 10 c.c." Will Occur	Pct. of Time that "1/5 in 10 c.c." Will Occur	Pct. of Time that "2/5 in 10 c.c." Will Occur
0	100	0	0
1	50	50	0
2	25	55	20
3	12.5	45	36
4	6.25	33.55	
5	3.13	23.26	
6	1.56	15.52	
7	0.78	10.09	
8	0.39	6.44	
9	0.19	4.05	
10	.....	2.54	
11	.....	1.57	

It will be noticed from these calculations that when 4 B. coli are present in the sample, the sample will pass the standard about 40

7. U. S. Pub. Health Rept., 1914, p. 2959.

percent of the time. And one out of about every six samples, containing 6 B. coli, will pass the standard (1.56 percent plus 15.52 percent). On the other hand, one out of every five samples containing only 2 B. coli per 100 c.c. will fail to pass the standard.

Consequently, when it is remembered that the standard signifies a most probable limit of 2 B. coli per 100 c.c., it is evident that the standard method of analysis renders the standard much more lenient than might, at first glance, be supposed.

(It is to be noticed that the most probable number corresponding to the result "1/5 in 10 c.c.", viz., 2, is the same as that given by the general formulae which do not allow for non-replacement. In general, the "most probable numbers" are quite the same, whether or not non-replacement is allowed for, so far as current methods, or systems, of analysis are concerned.)

METHODS FOR OBTAINING CONSIDERABLE PRECISION

In experimental work it is often necessary to know very approximately the number of fermenting organisms contained in the sample. Now, with the aid of the formulae given, the investigator may determine beforehand the details of his experiment in order that significant results may be obtained.

For instance, if two forms of apparatus are to be compared, it is evident that more than ten tubes must be used with each form; for two results such as "1/10" and "3/10" are of no significance as regards relative facility of growth with the two forms of apparatus; the curves of these results overlap so considerably that one would almost be justified in expecting just the opposite order of inference at the next experiment.

So also when even 100 tubes are used, the plotting of a few ordinates of the curves for the results will reveal the significance of these results and indicate whether or not the differences among these results may be called "significant."

If considerable precision is desired in the result to be obtained from examination of a water supply, the use of many tubes inoculated from one large sample is preferable to the use of a few tubes inoculated from each of several small samples. For, as has been shown, if the B. coli contents of the small samples vary, the average of the results obtained from these small samples is not trustworthy. But

the result from the single large sample (composite, if necessary) is trustworthy according to the number of tubes employed in its analysis.

Moreover, by plotting a few ordinates of the curve for the result the practical range of numbers of B. coli which may be in the sample is readily determined, in the case of the single large sample. But in the case of the several small samples, no such estimate of the precision of the average result is possible.

The only method of obtaining a good average result from several samples is, therefore, that of mixing the sample before the analysis, thus making a composite sample from which many tubes may be inoculated. And the precision which may be obtained in the result from this composite sample is limited only by the size of the sample.

VARIOUS METHODS OF MAKING DILUTIONS

By an extension of the mathematical analysis (too lengthy a development to be included here), it may be shown that the following methods of obtaining 0.1 c.c. of the sample allow the same probability of drawing some B. coli in that 0.1 c.c.: (1) taking, by means of a pipette, exactly 0.1 c.c. of the sample; (2) mixing exactly 1 c.c. of the sample with 9 c.c. of sterile water, and taking exactly 1 c.c. of the mixture; (3) mixing exactly 10 c.c. of sample with 90 c.c. of sterile water and taking exactly 1 c.c. of the mixture. In practice, however, the different facilities of error involved in measuring out the different quantities of sample and water, may upset this balance of probability.

EXPERIMENTAL

To test the fundamental general formulae given by theory, two sets of experiments were made.

With Seeds.—A certain number of white seeds (representing B. coli) were mixed with 100 volumes of black seeds (representing the sample). From this mixture were then drawn a certain number of volumes, and at each draw presence or absence of white seeds noted and recorded. These volumes were then replaced in the container, the whole thoroughly mixed again, and the first procedure repeated.

This was continued until 50, 100, or 200 trials had been made, as indicated in Table 3.

This procedure corresponds to the analysis of 50, 100, or 200 samples, each of which contained the same number of B. coli. It is to be noted that the conditions obtaining in the drawing of the several

volumes from each sample were exactly the same as those which obtain in actual analysis of a water sample: there was no replacement after each separate draw.

Table 3 shows the results obtained, together with the theoretical results which should have been obtained according to the general formulae of Case 2.

The agreement between experiment and theory is remarkably close, if the small number of trials is considered, despite the fact that the

TABLE 3

SHOWING THE NUMBER OF TIMES CERTAIN RESULTS OCCURRED IN THE EXPERIMENT, AND THE MOST PROBABLE NUMBER OF TIMES THESE RESULTS SHOULD HAVE OCCURRED ACCORDING TO THEORY

Result	40 White Seeds in 100 Volumes (200 Trials)		160 White Seeds in 100 Volumes (100 Trials)		160 White Seeds in 88 Volumes (150 Trials)*		240 White Seeds in 100 Volumes (50 Trials)		320 White Seeds in 100 Volumes (50 Trials)	
	Experiment	Theory	Experiment	Theory	Experiment	Theory	Experiment	Theory	Experiment	Theory
0/1	147	134	16	20	26	24	6	4	4	2
1/1	53	66	84	80	124	126	44	46	46	48
0/2	107	90	3	4	6	4	...	...	1	
1/2	80	89	26	32	49	41	9	8	6	4
2/2	13	21	71	64	95	105	41	42	43	46
0/5	41	27								
1/5	61	66	1	1	...	1				
2/5	59	65	2	5	5	5				
3/5	26	32	19	21	22	24	3	3	2	1
4/5	12	8	40	41	73	60	15	16	10	9
5/5	1	1	38	33	50	61	32	31	38	40
0/10	4	4								
1/10	21	18								
2/10	41	40								
3/10	55	52								
4/10	45	43	...	1	2					
5/10	23	27	2	3	2	2				
6/10	10	11	10	9	12	8	1			
7/10	...	3	16	20	34	23	3	2		
8/10	1	1	30	30	41	43	6	9	4	3
9/10	...	...	31	27	43	49	17	19	13	14
10/10	...	...	11	11	16	25	23	20	33	33

* The theoretical calculations for this case are based on "180 white seeds in 100 volumes," a condition which is roughly equivalent to that of "160 white seeds in 88 volumes." Theory gives practically identical probabilities for these two conditions.

It must be remembered that even in these experiments Chance is operating, and exact correspondence between Experiment and Theory is not to be expected with only one short series of trials. But with a greater and greater number of trials, the percentage "error," or difference between Experiment and Theory, may be expected to become correspondingly less.

theoretical results were calculated on the assumption of replacement after each separate draw. It is evident that, for all practical purposes, this assumption may be disregarded, and the general formulae applied directly to the fermentation-tube results which may be obtained in the laboratory.

With Bacillus coli.—A large bottle was almost filled with about 8,000 c.c. of normal salt solution, sterilized, and after it had been

cooled, carbon-dioxid was blown through it for a few minutes. (It was feared that practical absence of this gas might prove destructive to the organism to be added.)

A twenty-four-hour bile culture of B. coli was filtered through sterile filter-paper, and a few drops of the filtrate added to the solution, which was then well mixed and placed in the dark.

After twenty-four hours several bile tubes were inoculated with 1 c.c. of the solution, which, after another twenty-four hours, gave results indicating the number of B. coli in the sample to be such that no dilutions would be required for analysis by the fermentation test.

The sample after thorough mixing, was then tested as follows: With a 10 c.c. pipette graduated in tenths, 10 c.c. of the sample (8,000 c.c. in quantity) was withdrawn, and 1 c.c. inoculated into each of ten Durham tubes containing lactose peptone bile. The ten tubes were held together as a unit, being contained in a small zinc box into which they fitted.

After ten such boxes of tubes had been inoculated in this manner, the sample was well shaken, and another ten boxes of tubes inoculated; and so on, until 76 boxes of tubes, or 760 tubes, each contained 1 c.c. of the sample.

After forty-eight hours' incubation at 37 C., the number of positives were distributed as follows:

First 100 tubes	71 positives	
Next 100 tubes	79 positives	
Next 100 tubes	77 positives	
Next 100 tubes	79 positives	
Next 100 tubes	78 positives	
Next 100 tubes	85 positives	
Next 100 tubes	82 positives	
Next 60 tubes	46 positives	(77 per cent.)

Thus the total number of positives was 597* out of 760 inoculated. The most probable number of B. coli per 100 c.c. (capable of fermenting the media), in the sample, is then given by the solution of the equation:

$$1 - .99^{x} = \frac{597}{760}$$

and, this being solved, x=153, to the nearest unit.

Now, if the number of B. coli in the sample was really about 153 per 100 c.c., the number of times the particular results "10/10 in

* Every tube of the 597 positive tubes contained a large amount of gas, except one. This tube contained only about 2 percent of gas, but was recorded as positive. Seventy-two hours' incubation gave no further positives.

1 c.c." "9/10 in 1 c.c.", etc., occurred, should approximate the number of times these results should occur according to calculation by the general formulae of Case 2, with $x=153$. Thus, the result "10/10 in 1 c.c." should occur $76(1-.99^{153})^{10}=6.764$, or 7 times, to the nearest unit.

These calculations, together with the number of times the various results did actually occur, are given below:

Result	Experiment	Theory
4/10	0	0
5/10	4	3
6/10	7	8
7/10	15	17
8/10	27	23
9/10	16	18
10/10	7	7
	76	76

Taking only the results given by the first five tubes, the first two tubes, the first tube, of each box of ten tubes, we have the actual and calculated results as follows:

Result	Exp'm't	Theory	Result	Exp'm't	Theory	Result	Exp'm't	Theory
5/5	26	23	2/2	42	47	1/1	21	16
4/5	25	31	1/2	32	26	0/1	55	60
3/5	21	17	0/2	2	3			
2/5	4	5						
1/5	0	0						

The close correspondence between the experimental and theoretical results, afford an excellent verification of the "most probable number," 153 B. coli per 100 c.c., which was obtained by the other formula, and which was in no way dependent upon the formulae used in these latter calculations.

These experiments prove beyond doubt that the general formulae given in the early part of this paper are applicable to the general problem of numerical interpretation of fermentation-tube results. In all these experiments the procedure adopted was strictly analogous to that obtaining in laboratory practice; yet, despite the assumptions involved in the theoretical calculations, as regards quantity of sample and replacement after every draw, the agreement between experiment and theory is all that could be desired.

One particular point, which is especially emphasized by the results of these experiments, should be noted, namely, the persistence with which results of small probability (occurring with small frequency) do

occur. In nearly all the series of experimental results given in this paper, the number of results of small frequency check very closely the number indicated by theory. The conclusion is obvious: The odd and infrequent result is bound to occur with its allotted frequency.

SUMMARY

The frequency of the appearance of the fermenting organism in the volume drawn from the sample for the test is an exponential function of the number of such organisms in the sample.

Every fermentation-tube result, whether simple or compound, corresponds to one most probable number of organisms, and this number demands the same consideration as does the most probable number of bacteria corresponding to the plate-count in the estimation of other groups of bacteria. A simple method is available by which this most probable number of organisms may be calculated.

Odd results, such as "anomalies," are sure to occur with their theoretical frequency, and should not be thrown out of the record, but preserved and given their numerical interpretations.

Methods are available by which some knowledge of the degree of precision of a result may be obtained. In consequence, different results may be compared, the significance of the difference between certain results may be determined, and sampling may be so conducted as to give, in the result, any degree of precision desired.

Because of the difficulty of obtaining a trustworthy average of fermentation-tube results, the results should be expressed in full, together with a statement of the "most probable number" of organisms in the sample.

Calculation indicates that the precision of the fermentation-tube method compares not unfavorably with that of the plate method, of analysis.

The current methods of making dilutions should all give the same results, so far as chance distribution of the organisms is concerned.

Both calculation and experiment show that, for all practical purposes, the general formulae given are applicable to the problem of numerical interpretation of fermentation-tube results obtained in the laboratory.* For the severe conditions involved in the method of analysis prescribed by the U. S. Treasury Standard for waters sup-

* The writer has prepared a table which contains logarithms for the rapid calculation of "most probable numbers," values of the ordinates of the curves for the results N/1, N/2, N/3, N/5 and N/10, etc., copies of which he will be glad to furnish upon request. Address 9 St. James St., Montreal.

plied by common carriers, special formulae are necessary, which are given. (The sample should always be thoroughly mixed before testing. If not, the Probabilities, as herein developed, will not hold.)

For the proof that various methods of making dilutions give the same probability; for the methods of summing the ordinates of the curves; for the formula giving the probability of obtaining a specified number of organisms in the volume drawn, as well as for much helpful criticism and suggestion, the writer is greatly indebted to Professor Wm. D. Cairns, Associate Professor of Mathematics, Oberlin University, Ohio.

ADDITIONAL NOTE ON A METHOD OF EMPLOYING A LARGE NUMBER OF TUBES IN THE FERMENTATION TEST

It is evident from the foregoing study, that to attain even a fair degree of precision with the fermentation test, several tubes must be used at each dilution. But the use of several tubes is attended by considerable inconvenience when they are handled in the ordinary manner, and until this inconvenience is largely obviated by means of some improvement in apparatus, increase in precision in routine work will perhaps not become very general. For two years the Laboratory of the Board of Health of the Province of Quebec has employed the following system of analysis on all ordinary samples of water where the field survey does not reveal heavy pollution:

(1) Ten cubic centimeters of the sample are withdrawn by means of a straight-walled 10 c.c. pipette, and mixed with 90 c.c. of sterile water.

With another 10 c.c. pipette, 10 c.c. of this mixture are withdrawn and 1 c.c. delivered into each of ten Durham bile tubes. This gives ten tubes with 0.1 c.c. of the sample in each.

(2). With the other pipette, 10 c.c. of the sample are withdrawn and 1 c.c. delivered into each of ten tubes.

(3). With the same pipette 10 c.c. of the sample are delivered into each of two large tubes.

Results from this system will be of the form:

"A/2 in 10 c.c., B/10 in 1 c.c., C/10 in 0.1 c.c."

Such a procedure may appear laborious, but by means of a simple contrivance for handling the tubes, the labor involved is very little more than that in using only two tubes at each dilution in the ordinary manner. This contrivance consists simply of a small zinc box, into which ten Durham tubes fit rather loosely. The outer tube of the fermentation tube is a large, flat-bottomed specimen tube; the inner tube is of the same pattern but smaller.

When preparing the apparatus, the outer tubes are placed in position in the boxes, and filled to the proper height with the medium by moving from tube to tube a large cylindrical funnel provided with a rubber tube and pinch-cock. A large number of tubes may be filled very rapidly in this manner. The inner tubes are then dropped into place, and the cover of the box, into which has been placed a strip of cotton, is fitted over the tops of the tubes and secured by a rubber band, as shown in the accompanying figure. The one strip of cotton thus serves as a common plug for all the tubes, being held firmly in place by the pressure of the rubber band. If the tubes are not to be used at

once, they are wrapped in ordinary wrapping paper (which may be obtained of the desired width), before securing with the rubber band. The boxes are then sterilized in the ordinary way.

When using the tubes, the paper wrapper is removed, the cover lifted, and by means of a straight-walled 10 c.c. pipette, 10 c.c. of the sample are withdrawn and 1 c.c. of this delivered into each of the ten tubes in the box by simply moving the pipette from tube to tube. The cover is then dropped into

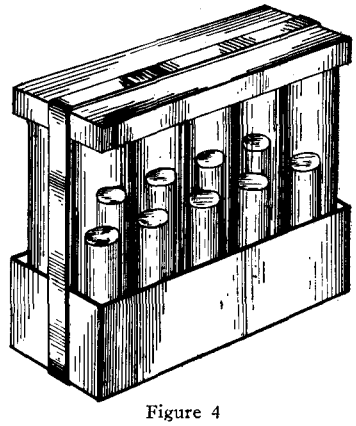

Figure 4

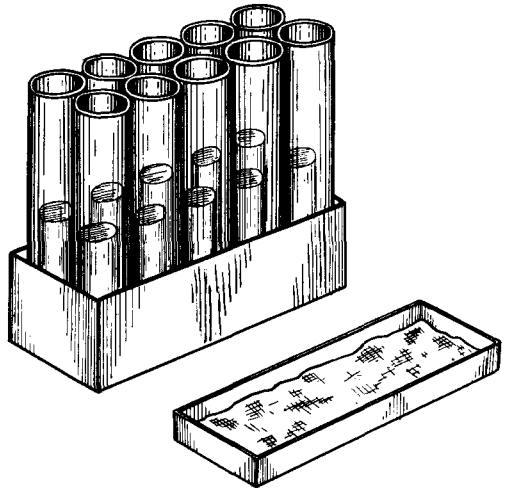

Figure 5

place, the rubber band slipped around the whole, and the box marked with the number of the sample and the number of the dilution. The box is then ready for the incubator, for re-wrapping with the paper is quite unnecessary.

The advantages of this method are many:

(1). A more reliable result (greater precision) is obtained with the large number of tubes, and with very little extra labor.

(2) All tubes at the same dilution for the same sample are held together as a unit, reading of the results being thus rendered a very simple operation.

(3) The box is marked, the inconvenience attending the marking of each individual tube being thus obviated.

(4). The shape of the boxes permits very easy handling, and very compact packing in the incubator, for they may be packed one upon the other.

(5) In experimental work, some such device is practically indispensable. For instance, in the experiment described in the foregoing paper in which 760 tubes were inoculated, each with 1 c.c. of the sample, it would have been quite out of the question to have attempted the inoculation of such a large number of tubes in the usual manner of handling each tube separately.

Two years' experience with this method of employing large numbers of tubes in the fermentation test have amply demonstrated the facility with which the apparatus may be prepared and manipulated, and the greater satisfaction to be derived from the corresponding increase in precision obtained by its use.